

Adaptive diversification within a large family of recently duplicated, placentally expressed genes

Austin L. Hughes^{*†}, Jonathan A. Green[‡], Juana M. Garbayo[‡], and R. Michael Roberts^{*§}

^{*}Department of Biology, Pennsylvania State University, University Park, PA 16808; and [‡]Departments of Animal Sciences and Biochemistry, University of Missouri, Columbia, MO 65211

Contributed by R. Michael Roberts, January 5, 2000

The pregnancy-associated glycoproteins (PAG) are putative peptide-binding proteins and products of a large family of genes whose expression is localized to the placental surface epithelium of artiodactyl species. We have tested the hypothesis that natural selection has favored diversification of these genes by examining patterns of nucleotide substitution in a sample of 28 closely related bovine, caprine, and ovine family members that are expressed only in trophoblast binucleate cells. Three observations were made. First, in codons encoding highly variable domains of the proteins, there was a greater accumulation of both synonymous and nonsynonymous mutations than in the more conserved regions of the genes. Second, in the variable regions, the mean number of nonsynonymous nucleotide substitutions per site was significantly greater than the mean number of synonymous substitutions per site. Third, nonsynonymous changes affecting amino acid charge occurred more frequently than expected under random substitution. This unusual pattern of nucleotide substitution implies that natural selection has acted to diversify these PAG molecules at the amino acid level, which in turn suggests that these molecules have undergone functional diversification. We estimate that the binucleate cell-expressed PAG originated 52 ± 6 million years ago, soon after the divergence of the ruminant lineage. Thus, rapid functional diversification of PAG expressed in trophoblast binucleate cells seems to have been associated with the origin of this unique placental adaptation.

aspartic proteinase | multigene families | positive selection | trophoblast | pregnancy-associated glycoprotein

The pregnancy-associated glycoproteins (PAGs) constitute an extensive family of recently duplicated genes within the aspartic proteinase superfamily. In cattle and sheep and probably also in other pecoran mammals, there may be as many as 100 or more genes (1) encoding these molecules. Thus far, 21 different cDNAs, differing by at least 5% in nucleotide sequence, have been cloned from the placenta of cattle (ref. 1 and J.A.G. and R.M.R., unpublished data), and another 11 have been cloned from sheep (1). PAGs have also been identified in pigs, a nonruminant species (2), and seem to occur throughout the Artiodactyla order (1). Only a single representative PAG-like gene has been found in the horse and zebra (order Perissodactyla; ref. 3), in the cat (Carnivora; ref. 4), and in the mouse (Rodentia; ref. 5), suggesting that the duplications that gave rise to the numerous ruminant PAG were initiated relatively recently and certainly since the separation of the Artiodactyla and Perissodactyla. In artiodactyls, the PAG are expressed in the outer epithelial cell layer of the placenta, which is in direct contact with the uterine endometrium. Their function at this interface between fetal and maternal tissue remains unknown; however, they are not believed to act primarily as proteinases, because many, possibly the majority, possess mutations that are likely to render them inactive as enzymes (6, 7). Structural modeling, however, suggests that they have retained the peptide-binding cleft of aspartic proteinases (7), and recent work has indicated that they are capable of binding short peptides (L. Landon, J.A.G.,

and R.M.R., unpublished data). The available PAG sequences from sheep and cow suggest that members of this gene family may have differentiated functionally (1).

The evolution of new protein function after gene duplication is believed to have played a major role in the evolution of life on earth (8, 9), but the mechanism by which this process occurs remains controversial. The most widely cited model for the evolution of new protein function assumes that new function arises as a result of selectively neutral mutations that are fixed by chance in a redundant gene copy and fortuitously provide it with a new function (10). However, molecular studies have provided several lines of evidence suggesting that new gene functions rarely, if ever, evolve in this way (9, 11). For example, several studies have shown that when genes differentiate functionally, gene duplication is often followed by positive Darwinian selection acting on functionally important regions of proteins. Positive selection is revealed by comparing the patterns of synonymous and nonsynonymous (amino acid-altering) nucleotide substitution. In most genes, the number of synonymous substitutions per synonymous site exceeds the number of nonsynonymous substitutions per nonsynonymous site (12). This pattern occurs, because most amino acid changes are selectively disadvantageous and thus are removed by natural selection (so-called "purifying" selection; ref. 13). The opposite pattern is evidence that natural selection has acted to favor amino acid changes (9). Such evidence has been obtained for comparisons between recently duplicated members of a number of multigene families (e.g., refs. 9 and 14–17). The present paper describes the use of this approach to test for evidence of adaptive diversification in the PAG gene family of Bovidae.

Methods

DNA Sequences Analyzed. Bovine PAG cDNAs were isolated from cDNA libraries prepared at day 19 (whole concepti), day 25 (dissected trophoblast; ref. 1), and day 260 of pregnancy (whole placental tissue; ref. 6) as described elsewhere (see ref. 1). Additional bovine transcripts were cloned by reverse transcription-PCR procedures from RNA isolated at day 19 (1), at day 25, and at term.

The primers used were a boPAG1 5' oligonucleotide (5'-AGGAAAGAAGCATGAAGTGGCT3'-, with the start codon underlined; ref. 6), whose sequence is completely or

Abbreviation: PAG, pregnancy-associated glycoprotein.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. M73961, L06153, AF020506–AF020514, AF192330–AF192338, AF191326–AF191336, M73962, U30251, and U94789–U94793).

[†]Present Address: Department of Biological Sciences, University of South Carolina, Columbia, SC 29208.

[§]To whom reprint requests should be addressed. E-mail: robertsrsm@missouri.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.050002797. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.050002797

```

BOPAG1 MKWLVLGLV AFSECTVIRP LRLKTMNRV VSGKNMLNMF LKEHAYSLSQ ISFSGSNLTT
BOPAG5 .....L TS.....IL .TKV....KT L.E.....Q.R.....S.....I.I
BOPAG6 .....V.....V.....A I.....T.....I .....R.P.....
BOPAG15 .....V.....V.....TKT L.....T.....R.....H.....I
BOPAG16 .....V.....KT L.....T.....P.R.....H.....
BOPAG20 .....F.....V.....KT L.....P.K.....

* ***** *
BOPAG1 HPLRNKIDLV YMGNITIGTP PQEFQVVDFT ASSDLWVPSD FCTSPACSTH VFRHLQSSST
BOPAG5 .....M.M.....V.K.....G.....E.....V.....P.S.....I.....E.....
BOPAG6 .....R.....F.....V.....L.....G.....A.....I.....N.SS.AA.....H.....
BOPAG15 .....R.....F.....L.....G.....K.....F.....
BOPAG16 L.....R.ML.....V.....G.....I.....F.....
BOPAG20 L.....W.IF.....I.T.....I.....I.....N.ST.....R.....

*****
BOPAG1 FRLTNKTFRI TYGSGRMKG VVHDTVRIGN LVSTDQPFGL SIEEYGFEGR IYDGVGLNLY
BOPAG5 SG.....Q.....S.....ST.F LAY.....D L.....E.....M.H.....DL PF.....
BOPAG6 .....P.....D.....CLKDS.K.I PF.....I.....S.....
BOPAG15 .....S.....E.....T.E.I A.....D .....MT.S.I PF.....
BOPAG16 .....P.T.....I.....A.....MA.....L.S RF.....I.....
BOPAG20 .....G.....D.....VA.....RF.....

*****
BOPAG1 PNISFSGAIP IFDKLKNQRA ISEPVFAFY L SKDEREGSVV MFGVDHRY Y EGELNWWPLI
BOPAG5 .DM..ITT.....N.....G F.....G.VK.....T.....K.....
BOPAG6 .....KT.....F.....K.....K.....
BOPAG15 .....G.....Q.....K.....K.....
BOPAG16 .....L.C.....G.....D.I.....K.....K.....
BOPAG20 .....K.....EG.....KQK.....K.....

*****
BOPAG1 QAGDWSVHMD RISTIERKIITA CSDGCKALVD TGTSDIVGPR RLWNHRII GAIPRGSEHY
BOPAG5 .....E.L.....MK.V.....G.E.FY.....L.L.....QK.....T.Q.....
BOPAG6 .....V.....F.....TMK.V.....S.....T.....WK.....R.R.L.PQYF
BOPAG15 .....I.....C.MR.V.....G.E.V.....V.M.K.K.T.D.QK.....TL.FK.....
BOPAG16 R.....I.V.....TMK.EV.....A.....L.Q.G.VID.....K.....T.....K.....
BOPAG20 R.....V.....TMKGEV.G.....T.M.....S.N.Q.G.VID.....K.....T.....K.....

*****
BOPAG1 VPCSEVNTLP SIVFTINGIN YPVGGRAYIL KDRGRCYTT FQENRVSSST ETWYLGDFVL
BOPAG5 IS.FA.IS.....I.....I.....A.....H.....S.....H.....P.....K.T.T.....I.....
BOPAG6 .....S.A.....I.....RL.A.....H.....S.....A.....K.H.F.PI.....L.....
BOPAG15 .....S.A.D.....T.....R.A.....S.....C.SS.....TT.P.....I.....
BOPAG16 .....S.V.....I.....AP.....S.....Y.....A.....K.Q.....RR.....S.L.....
BOPAG20 .....S.A.SA.....V.....A.....V.....PT.N.....K.K.....RR.....F.I.....EA.....

*****
BOPAG1 RLYFSVFDGR NDRIGLARAV
BOPAG5 .....Q.....
BOPAG6 .....R.....
BOPAG15 .....
BOPAG16 .....
BOPAG20 .....

```

Fig. 1. Alignment of selected PAG. Variable regions are marked with asterisks.

almost completely conserved in all the PAGs cloned thus far, and one of the following 3' oligonucleotides: 5'-GCGCTC-GAGTTACACTGCCCGTGCCAGGC-3', 5'-GTTCTC-GAGCTTACACAGCAGGAGCCAAGCC-3', 5'-GCGCTC-GAGTTGAAGCAGCTCCAGCATTTA-3', and 5'-GT-TCTCGAGCTTATACCGCAGTAGCCAGTCC-3'. The latter four oligonucleotides represented sequences encompassing the end of the ORFs of known PAG genes (the stop codon is underlined). The reverse transcription-PCR products were cloned into the pGEM-T Easy vector (Promega). The positive clones were isolated, analyzed for size by restriction mapping, and partially sequenced. Those cDNAs sharing less than 95% identity with other bovine PAG were fully sequenced in both directions.

The cloning of the ovine PAG transcripts has been described in detail elsewhere (1).

To clone caprine PAG transcripts, RNA from day 45, 65, 87, and 115 placental tissue was reverse transcribed by using the same primers in a manner identical to that described above for boPAG transcripts.

Statistical Methods. PAG sequences were aligned at the amino acid level by using the CLUSTAL W program (18). Fig. 1 shows a partial alignment; the complete alignment is available from the authors on request. In any set of pairwise comparisons among sequences, any amino acid position at which the alignment postulated a gap was excluded from all pairwise comparisons so that the same set of positions was compared in each case. By

inspection of the aligned amino acid sequences, we identified four highly variable regions (Fig. 1). The third of these regions includes the B lobe catalytic domain of aspartic proteinases and its conserved signature motif (ref. 1; residues 268–273 in Fig. 1). The fourth region includes an invariant phenylalanine residue (position 341 in Fig. 1), which is involved in bridging between two loops (1).

Phylogenetic trees were constructed by means of the maximum parsimony method (19) and the neighbor-joining method (20). Neighbor-joining trees were constructed on the basis of the following distances: the uncorrected proportion of amino acid difference (p), the γ -corrected estimate of the number of amino acid replacements per site (21), and the number of nonsynonymous nucleotide substitutions per site (d_N ; ref. 22). Synonymous sites were not used because synonymous sites were saturated with changes in certain comparisons among these sequences. The reliability of branching patterns in phylogenetic trees was tested by bootstrapping, which involved repeated pseudosampling from the data (with replacement; ref. 23); 1,000 bootstrap pseudosamples were used. All of these methods yielded essentially the same results; therefore, only the neighbor-joining tree based on p is presented here.

Numbers of synonymous (d_S) and nonsynonymous (d_N) nucleotide substitutions per site were estimated by Nei and Gojori's method (22), modified as recommended by Zhang *et al.* (17) to take into account different frequencies of transitions and transversions. The original method (22) assumes equal usage of all four nucleotides in counting fractional synonymous and nonsynonymous sites. However, this method will lead to an underestimation of the number of synonymous sites, particularly in the case of 2-fold degenerate sites at which transitions are synonymous and transversions are nonsynonymous (23). The observed frequency of transitions and transversions at 2-fold degenerate sites is expected to reflect the effects both of mutational bias and of purifying selection, which is expected to eliminate many synonymous mutations. Therefore, it is inappropriate to estimate the transition:transversion ratio from such sites. Because all mutations at 4-fold degenerate sites are synonymous, these sites are most appropriate for estimating the transition:transversion ratio. Therefore, we estimated the transition:transversion ratio from 4-fold degenerate sites.

Hughes *et al.* (24) developed a method of testing the hypothesis that nonsynonymous differences occur in such a way as to change some amino acid property of interest to a greater extent than is expected under random substitution. This method involves computing the proportion of conservative nonsynonymous difference per conservative nonsynonymous site (p_{NC}) and the proportion of radical nonsynonymous difference per radical nonsynonymous site (p_{NR}). When $p_{NC} > p_{NR}$, nonsynonymous differences occur in such a way as to conserve the property of interest. When $p_{NC} = p_{NR}$, nonsynonymous differences occur at random with respect to the property of interest. When $p_{NR} > p_{NC}$, nonsynonymous differences occur in such a way as to change the property of interest to a greater extent than expected at random. Because this last pattern indicates a directional change in amino acid sequence, it is suggestive of positive selection (24).

For use in the present analyses, we developed a modified version of the earlier method (24) that takes into account transition:transversion ratio. Taking transitional bias into account is important in such analyses, because transitional mutations in nonsynonymous sites are often conservative with respect to amino acid properties such as charge or polarity, whereas transversal mutations are frequently radical. (A computer program for this method is available from A.L.H. on request.) We applied this method with respect to two amino acid properties, charge and polarity. With respect to charge, amino acid residues were categorized as positive (H, K, and R), negative (D

Table 2. Mean percentages of conservative (p_{NC}) and radical (p_{NR}) nonsynonymous differences per site, with respect to amino acid residue charge and polarity in variable regions and remainders of PAG genes

Property	p_{NC} , %	p_{NR} , %
Charge		
Variable	20.6 ± 1.5	30.3 ± 2.6**
Remainder	10.0 ± 0.6	11.2 ± 0.9
Polarity		
Variable	24.5 ± 1.7	23.0 ± 2.3
Remainder	10.9 ± 0.6	9.2 ± 0.8

Values are shown as mean percentages ± SEM. Tests of the hypothesis that $p_{NC} = p_{NR}$: **, $P < 0.01$.

Results

In the phylogenetic tree (Fig. 2), sequences from bovine, sheep, and goat PAGs are intermixed, indicating that many gene duplications in this family took place before these three species diverged. The phylogeny enabled us to identify a subgroup of closely related sequences from the three species (marked by bracket in Fig. 2) in comparisons among which synonymous sites were well below saturation. We analyzed this group of 28 sequences in testing for positive selection.

In pairwise comparisons among these 28 closely related sequences, mean d_N was significantly greater than mean d_S in the variable regions (Table 1). In the remainder of the gene, mean d_S was higher than mean d_N , although the difference was not statistically significant (Table 1). Both mean d_S and mean d_N in the variable regions were significantly greater than the corresponding values for the remainder of the gene (Table 1).

The pattern of synonymous and nonsynonymous nucleotide substitution is illustrated further in Fig. 3, where d_N is plotted against d_S for each pairwise comparison. In the variable regions, d_N exceeded d_S in 92.3% of pairwise comparisons (349 of 378), as indicated by points above the 45° line in Fig. 3A. The value d_N was particularly likely to exceed d_S , when d_S was relatively low (Fig. 3A). By contrast, in the remainder of PAG genes, d_S exceeded d_N in 74.1% (280 of 378) of pairwise comparisons (Fig. 3B).

When p_{NC} and p_{NR} were estimated with respect to amino acid residue charge, p_{NR} exceeded p_{NC} in the variable regions (Table 2). Thus, in the variable regions, amino acid changes occur disproportionately in such a way as to cause residue charge change. By contrast, in the remainder of PAG genes, p_{NC} and p_{NR} were not significantly different with respect to residue charge (Table 2). In neither the variable regions nor the remainder of PAG genes was there a significant difference between p_{NC} and p_{NR} with respect to residue polarity (Table 2).

Mean d_S values were used to estimate the times of origin of ruminant PAG genes. To obtain a calibration, apparently orthologous pairs of genes were compared between cattle (subfamily Bovine) and sheep and goats (subfamily Caprinae), as indicated in the phylogenetic tree in Fig. 2. Comparisons were made between boPAG8 and caPAG8 and between boPAG11, caPAG2, and ovPAG2. Mean d_S for these comparisons was 0.082 ± 0.016 . Assuming that the two lineages diverged 20 million years ago (26), the divergence time between the 28 trophoblast binucleate cell-expressed genes and the nearest out group (ovPAG5) was estimated at 52 ± 6 million years ago (mean $d_S = 0.214 \pm 0.023$). By using the same calibration, the deepest branch point in the phylogeny of Fig. 2 was estimated at 87 ± 6 million years ago (mean $d_S = 0.359 \pm 0.023$).

Discussion

These results indicate that positive Darwinian selection has acted to promote diversity at the amino acid level in the variable

regions of duplicated PAG genes. Furthermore, this selection has acted to promote amino acid residue charge changes to a greater extent than expected under random substitution. The results thus suggest that the variable regions are important for functional specificity of PAG proteins and that the pattern of amino acid residue charges in these regions may play a particularly important role in functional differences among members of this family. These variable regions represent surface-exposed loops (1), and we have suggested earlier that small additive changes in the packing of these loop regions influence ligand binding within the substrate binding clefts, thereby providing PAGs with a considerable potential range of peptide-binding specificities (1).

The pattern of synonymous and nonsynonymous substitution seen in the variable regions (Fig. 3A), where d_N was particularly likely to exceed d_S when the latter was low, seems to be characteristic of positive selection leading to diversification of genes within multigene families (9, 15, 16). This pattern evidently occurs, because there is a burst of positively selected nonsynonymous substitutions immediately after gene duplication, as the duplicate genes adapt to distinct functions. Because d_S is relatively low in comparisons between closely related sequences, this burst of positively selected nonsynonymous substitutions is likely to cause d_N to exceed d_S in such comparisons. However, once each duplicate gene has adapted to its specific function, purifying selection is expected to predominate, allowing the number of synonymous substitutions per site to catch up to and eventually exceed the number of nonsynonymous substitutions per site (9). As a result, it is generally possible to detect such selection only within a certain relatively limited time after gene duplication, probably 30–50 million years (at most) in the case of eukaryotes (9). Because genes of the PAG family have duplicated repeatedly within the order Artiodactyla, they provide a family of relatively recently duplicated and functionally differentiated genes ideal for studying the process of adaptive diversification of duplicate genes.

One surprising result was the finding that mean d_S in the variable regions was significantly higher than the corresponding value for the remainder of the genes (Table 1). Because d_S is expected to reflect the mutation rate, it is predicted that d_S will be relatively constant across different regions of a protein-coding gene (27). In most cases, analyses have supported this prediction (9). Exceptions to the rule can occur in the following circumstances: when the mutation rate varies over the length of a gene, as has sometimes been observed, especially in genes with long coding regions such as the cystic fibrosis transmembrane transporter genes (28), or alternatively, when different portions of two related genes differ in divergence time as a result of interlocus recombination or “gene conversion” (29). Given that the variable regions analyzed herein consist of four short isolated gene regions, it seems unlikely that these regions have higher mutation rates than the rest of the gene. Thus, it seems most plausible that d_S in the variable region genes exceeds that in the remainder of the genes, because interlocus gene conversion events have caused some degree of homogenization of loci outside the variable regions.

It is interesting that the 28 closely related PAG genes analyzed in Fig. 2 are expressed predominantly in trophoblast binucleate cells (ref. 1 and J.A.G., J.M.G., and R.M.R., unpublished data). These large cells, which comprise 15–20% of postattachment trophoctoderm (30), can migrate from the trophoblast and fuse with uterine epithelial cells to form either a syncytium (as in the sheep, in the goat, and in the early stages of placenta of the cow) or short-lived trinucleated cells (in established cow placenta; refs. 30 and 31). These cells constitute a unique feature of the synepitheliochorial placenta of the ruminant artiodactyl species. This type of placenta, with very limited invasive poten-

tial, probably developed from the noninvasive epitheliochorial placenta observed in nonruminant artiodactyls, such as camels (32) and pigs (33). The calculation of 52 ± 6 million years for the origin of the binucleate cell-specific PAGs is only slightly lower than recent estimates for the divergence of the Ruminantia and the Suidae (26). It is tempting to speculate that the burst of duplications that gave rise to the binucleate cell PAG group is linked to the emergence of this placental type within Ruminantia. Similarly, the origin of the Artiodactyla has been estimated

at about 83 million years ago (26), a value that is very close to our estimate of when the PAG genes as a whole first began to duplicate.

We thank Dr. Joseph M. Quattro for carefully reading the manuscript and for providing comments. This work was supported by National Institutes of Health Grants GM34940 and HD35898 to A.L.H. and R.M.R., respectively, and Grant 96-35205-3257 from the U.S. Department of Agriculture Competitive Grants Program to R.M.R.

1. Xie, S., Green, J., Bixby, J. B., Szafranska, B., DeMartini, J. C., Hecht, S. & Roberts, R. M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12809–12816.
2. Szafranska, B., Xie, S., Green, J. & Roberts, R. M. (1995) *Biol. Reprod.* **53**, 21–28.
3. Green, J. A., Xie, S., Szafranska, B., Gan, X., Newman, A. G., McDowell, K. & Roberts, R. M. (1999) *Biol. Reprod.* **60**, 1069–1077.
4. Gan, X., Xie, S., Green, J. & Roberts, R. M. (1997) *Biol. Reprod.* **56**, Suppl. 1, 191 (abstr.).
5. Chen, X., Roberts, R. M. & Green, J. A. (1999) *Biol. Reprod.* **60**, Suppl. 1, 216 (abstr.).
6. Xie, S. C., Low, B. G., Nagel, R. J., Kramer, K. K., Anthony, R. V., Zoli, A. P., Beckers, J. F. & Roberts, R. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10247–10251.
7. Guruprasad, K., Blundell, T. L., Xie, S., Green, J., Szafranska, B., Nagel, R. J., McDowell, K., Baker, C. B. & Roberts, R. M. (1996) *Protein Eng.* **9**, 849–856.
8. Li, W.-H. (1983) in *Evolution of Genes and Proteins*, eds. Nei, M. & Koehn, R. K. (Sinauer, Sunderland, MA), pp. 14–37.
9. Hughes, A. L. (1999) *Adaptive Evolution of Genes and Genomes* (Oxford Univ. Press, New York).
10. Ohno, S. (1973) *Nature (London)* **244**, 259–262.
11. Hughes, A. L. (1994) *Proc. R. Soc. London Ser. B* **23**, 119–124.
12. Li, W.-H. (1997) *Molecular Evolution* (Sinauer, Sunderland MA).
13. Kimura, M. (1977) *Nature (London)* **267**, 275–276.
14. Hill, R. E. & Hastie, W. D. (1987) *Nature (London)* **326**, 96–99.
15. Tanaka, T. & Nei, M. (1989) *Mol. Biol. Evol.* **6**, 447–459.
16. Hughes, A. L. & Yeager, M. (1997) *J. Mol. Evol.* **44**, 675–682.
17. Zhang, J., Rosenberg, H. F. & Nei, M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 3708–3713.
18. Thompson, J. D., Higgins, D. G. & Gibson, J. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680.
19. Swofford, D. L. (1990) PAUP, Phylogenetic Analysis Using Parsimony (Illinois Natural History Survey, Champaign, IL), Version 3.0.
20. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
21. Ota, T. & Nei, M. (1994) *J. Mol. Evol.* **38**, 642–643.
22. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3**, 418–426.
23. Li, W.-H. (1993) *J. Mol. Evol.* **36**, 96–99.
24. Hughes, A. L., Ota, T. & Nei, M. (1990) *Mol. Biol. Evol.* **7**, 515–524.
25. Nei, M. & Jin, L. (1989) *Mol. Biol. Evol.* **4**, 406–425.
26. Kumar, S. & Hedges, S. B. (1998) *Nature (London)* **392**, 917–920.
27. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335**, 167–170.
28. Hughes, A. L. (1994) *Mol. Biol. Evol.* **11**, 899–910.
29. Ohta, T. (1992) *Mol. Phyl. Evol.* **1**, 305–311.
30. Wooding, F. B. (1992) *Placenta* **13**, 101–113.
31. Hoffman, L. H. & Wooding, F. B. P. (1993) *J. Exp. Zool.* **266**, 559–577.
32. Skidmore, J. A., Wooding, F. B. P. & Allen, W. R. (1996) *Placenta* **17**, 253–262.
33. Friess, A. E., Sinowatz, F., Skolek-Winnisch, R. & Trautner, W. (1980) *Anat. Embryol.* **158**, 179–191.